Help Me Understand Genetics

Genomic Research

Reprinted from MedlinePlus Genetics

U.S. National Library of Medicine

National Institutes of Health

Department of Health & Human Services

Table of Contents

Genomic Research

1 What are single nucleotide polymorphisms (SNPs)?

Single nucleotide polymorphisms, frequently called SNPs (pronounced “snips”), are the most common type of genetic variation among people. Each SNP represents a difference in a single DNA building block, called a nucleotide. For example, a SNP may replace the nucleotide cytosine (C) with the nucleotide thymine (T) in a certain stretch of DNA.

SNPs occur normally throughout a person’s DNA. They occur almost once in every 1, 000 nucleotides on average, which means there are roughly 4 to 5 million SNPs in a person's genome. These variations occur in many individuals; to be classified as a SNP, a variant is found in at least 1 percent of the population. Scientists have found more than 600 million SNPs in populations around the world.

SNPs differ from substitution variants, which replace one DNA building block (nucleotide) with another. Substitution variants usually cause disease and are generally not found in 1 percent of any population. Additionally, SNPs differ from copy number variants (CNVs) , which occur when a whole gene (or other large section of DNA) is duplicated or deleted. Most commonly, SNPs are found in the DNA between genes. They can act as biological markers, helping scientists locate genes that are associated with disease. When SNPs occur within a gene or in a regulatory region near a gene, they may play a more direct role in disease by affecting the gene’s function.

Most SNPs have no effect on health or development. Some of these genetic differences, however, have proven to be very important in the study of human health. SNPs help predict an individual’s response to certain drugs, susceptibility to environmental factors such as toxins, and risk of developing diseases. SNPs can also be used to track the inheritance of disease-associated genetic variants within families. Research is ongoing to identify SNPs associated with complex diseases such as heart disease, diabetes, and cancer.

For more information about SNPs:

An audio definition of SNPs (https://www.genome.gov/genetics-glossary/Single-Nucleotide-Polymorphisms) is available from the National Human Genome Research Institute’s Talking Glossary of Genetic Terms.

How scientists locate SNPs in the genome (https://learn.genetics.utah.edu/content/precision/snips/) is explained by the University of Utah Genetic Science Learning Center.

For people interested in more technical data, the National Library of Medicine maintains a frequently updated resource at NCBI of a database of single nucleotide polymorphisms (dbSNP) (https://www.ncbi.nlm.nih.gov/SNP/)

2 What are genome-wide association studies?

Genome-wide association studies (GWAS) help scientists identify genes associated with a particular disease (or another trait). This method studies the entire set of DNA (the genome) of a large group of people, searching for small variations, called single nucleotide polymorphisms or SNPs (pronounced “snips”). Each study can look at hundreds or thousands of SNPs at the same time. Scientists can then identify SNPs that occur more frequently in people with a certain disease than in people without it. These SNPs are said to be associated with the disease, and they can help researchers pinpoint genes that are likely involved in disease development. Because genome-wide association studies examine SNPs across the genome, they represent a promising way to study complex, common diseases in which many genetic variations contribute to a person’s risk. This approach has identified SNPs associated with several complex conditions including diabetes, heart disease, Parkinson disease, and Crohn disease. SNPs have also been associated with a person’s response to certain drugs and susceptibility to certain environmental factors such as toxins. Researchers hope that future genome-wide association studies will identify additional SNPs associated with chronic diseases and drug effects.

Through genome-wide association studies, individual SNPs are identified that account for only a small percentage of disease risk. Together, large numbers of SNPs across the genome can help determine the overall risk of developing a disease or responding to particular drugs. Researchers can use information learned from genome-wide association studies to predict more accurately which prevention and treatment strategies will work in which groups of people, an important step in precision medicine.

For more information about genome-wide association studies:

The National Human Genome Research Institute provides a detailed explanation of genome-wide association studies (https://www.genome.gov/about-genomics/fact-sheets/Genome-Wide-Association-Studies-Fact-Sheet). In addition, the National Human Genome Research Institute and the European Bioinformatics Institute jointly provide a Catalog of Published Genome-Wide Association Studies (https://www.ebi.ac.uk/gwas/).

ClinicalTrials.gov (https://clinicaltrials.gov/), a service of the National Institutes of Health, provides easy access to information about clinical trials. This site offers the option to search for a specific clinical trial or browse by health condition or sponsor. They also provide a list of genome-wide association studies (https://clinicaltrials.gov/search?term=GWAS+OR+%22Genome+Wide+Association%22) that are accepting (or will accept) participants.

For more technical information, the NCBI’s Database of Genotypes and Phenotypes (dbGaP) (https://www.ncbi.nlm.nih.gov/gap/) contains data from genome-wide association studies.

3 What is pharmacogenomics?

Pharmacogenomics is the study of how genes affect a person's response to drugs. This field combines pharmacology (the science of drugs) and genomics (the study of genes and their functions) to develop effective, safe medications that can be prescribed based on a person's genetic makeup.

Many drugs that are currently available are "one size fits all," but they don't work the same way for everyone. It can be difficult to predict who will benefit from a medication, who will not respond at all, and who will experience negative side effects (called adverse drug reactions). Adverse drug reactions are a significant cause of hospitalizations and deaths in the United States.

Researchers are learning how variants in genes affect the body's response to medications. These genetic differences will be used to predict whether a medication will be effective for a particular person and which dose will help prevent adverse drug reactions. Conditions that affect a person's response to certain drugs include clopidogrel resistance, warfarin sensitivity, warfarin resistance, malignant hyperthermia, Stevens-Johnson syndrome/toxic epidermal necrolysis, and thiopurine S-methyltransferase deficiency.

The field of pharmacogenomics is growing, and new approaches are under study in clinical trials. In the future, pharmacogenomics will be used to develop tailored drugs to treat a wide range of health problems, including cardiovascular disease, Alzheimer disease, cancer, and asthma.

For more information about pharmacogenomics:

MedlinePlus provides additional details about pharmacogenetic tests (https://medlineplus.gov/lab-tests/pharmacogenetic-tests/).

The National Institute of General Medical Sciences offers a list of Frequently Asked Questions about Pharmacogenomics (https://www.nigms.nih.gov/education/fact-sheets/Pages/pharmacogenomics.aspx).

A pharmacogenomics fact sheet (https://www.genome.gov/dna-day/15-ways/pharmacogenomics) and a list of Frequently Asked Questions about Pharmacogenomics (https://www.genome.gov/FAQ/Pharmacogenomics) is offered by the National Human Genome Research Institute.

Medical Genetics Summaries (https://www.ncbi.nlm.nih.gov/books/NBK61999/), provided by the National Center for Biotechnology Information at the National Library of Medicine, provides information about specific genetic variants and how they can impact drug responses.

Additional information about pharmacogenetics is available from the Centre for Genetics Education (https://www.genetics.edu.au/SitePages/Pharmacogenomics.aspx) as well as Genes In Life (http://www.genesinlife.org/testing-services/testing-genetic-conditions/pha

rmacogenomic-testing).

PharmGKB (https://www.pharmgkb.org/) is a pharmacogenomics resource sponsored by the National Institutes of Health that collects information on human genetic variation and drug responses.

A list of clinical trials involving pharmacogenomics (https://clinicaltrials.gov/search/term=pharmacogenomics+OR+pharmacogenetics) is available from ClinicalTrials.gov, a service of the National Institutes of Health.

4 What are genome editing and CRISPR-Cas9?

Genome editing (also called gene editing) is a group of technologies that give scientists the ability to change an organism's DNA. These technologies allow genetic material to be added, removed, or altered at particular locations in the genome. Several approaches to genome editing have been developed. A well-known one is called CRISPR-Cas9, which is short for clustered regularly interspaced short palindromic repeats and CRISPR-associated protein 9. The CRISPR-Cas9 system has generated a lot of excitement in the scientific community because it is faster, cheaper, more accurate, and more efficient than other genome editing methods.

CRISPR-Cas9 was adapted from a naturally occurring genome editing system that bacteria use as an immune defense. When infected with viruses, bacteria capture small pieces of the viruses' DNA and insert them into their own DNA in a particular pattern to create segments known as CRISPR arrays. The CRISPR arrays allow the bacteria to " remember" the viruses (or closely related ones). If the viruses attack again, the bacteria produce RNA segments from the CRISPR arrays that recognize and attach to specific regions of the viruses' DNA. The bacteria then use Cas9 or a similar enzyme to cut the DNA apart, which disables the virus.

Researchers adapted this immune defense system to edit DNA. They create a small piece of RNA with a short "guide" sequence that attaches (binds) to a specific target sequence in a cell's DNA, much like the RNA segments bacteria produce from the CRISPR array. This guide RNA also attaches to the Cas9 enzyme. When introduced into cells, the guide RNA recognizes the intended DNA sequence, and the Cas9 enzyme cuts the DNA at the targeted location, mirroring the process in bacteria. Although Cas9 is the enzyme that is used most often, other enzymes (for example Cpf1) can also be used. Once the DNA is cut, researchers use the cell's own DNA repair machinery to add or delete pieces of genetic material, or to make changes to the DNA by replacing an existing segment with a customized DNA sequence.

Genome editing is of great interest in the prevention and treatment of human diseases. Currently, genome editing is used in cells and animal models in research labs to understand diseases. Scientists are still working to determine whether this approach is safe and effective for use in people. It is being explored in research and clinical trials for a wide variety of diseases, including single-gene disorders such as cystic fibrosis, hemophilia, and sickle cell disease. It also holds promise for the treatment and prevention of more complex diseases, such as cancer, heart disease, mental illness, and human immunodeficiency virus (HIV) infection.

Ethical concerns arise when genome editing, using technologies such as CRISPR-Cas9, is used to alter human genomes. Most of the changes introduced with genome editing are limited to somatic cells, which are cells other than egg and sperm cells (germline cells). These changes are isolated to only certain tissues and are not passed from one generation to the next. However, changes made to genes in egg or sperm cells or to the genes of an embryo could be passed to future generations. Germline cell and embryo genome editing bring up a number of ethical challenges, including whether it would be permissible to use this technology to enhance normal human traits (such as height or

intelligence). Based on concerns about ethics and safety, germline cell and embryo genome editing are currently illegal in the United States and many other countries.

Scientific journal articles for further reading

Ormond KE(1), Mortlock DP(2), Scholes DT(3), Bombard Y(4), Brody LC(5), Faucett WA(6), Garrison NA(7), Hercher L(8), Isasi R(9), Middleton A(10), Musunuru K(11), Shriner D(12), Virani A(13), Young CE(3). Human Germline Genome Editing. Am J Hum Genet. 2017 Aug 3;101(2):167-176. PubMed: 28777929. Free full-text available from PubMed Central: PMC5544380.

Gupta RM, Musunuru K. Expanding the genetic editing tool kit: ZFNs, TALENs, and CRISPR-Cas9. J Clin Invest. 2014 Oct;124(10):4154-61. doi: 10.1172/JCI72992. Epub 2014 Oct 1. Review. PubMed: 25271723. Free full-text available from PubMed Central: PMC4191047.

Hsu PD, Lander ES, Zhang F. Development and applications of CRISPR-Cas9 for genome engineering. Cell. 2014 Jun 5;157(6):1262-78. doi:10.1016/j.cell.2014.05.010. Review. PubMed: 24906146. Free full-text available from PubMed Central: PMC4343198.

Komor AC, Badran AH, Liu DR. CRISPR-Based Technologies for the Manipulation of Eukaryotic Genomes. Cell. 2017 Apr 20;169(3):559. doi:10.1016/j.cell.2017.04.005. PubMed: 28431253.

For more information about CRISPR-Cas9 and other genome editing technologies:

The National Human Genome Research Institute has a series of fact sheets about genome editing:

- Overview of genome editing (https://www.genome.gov/about-genomics/policy-issues/what-is-Genome-Editing)
- How does genome editing work? (https://www.genome.gov/about-genomics/policy-issues/Genome-Editing/How-genome-editing-works)
- How is genome editing used? (https://www.genome.gov/about-genomics/policy-issues/Genome-Editing/How-genome-editing-is-used)
- What are the ethical concerns about genome editing? (https://www.genome.gov/about-genomics/policy-issues/Genome-Editing/ethical-concerns)
- What do people think about genome editing? (https://www.genome.gov/about-genomics/policy-issues/Genome-Editing/public-opinion)
- What's happening in genome editing right now? (https://www.genome.gov/about-genomics/policy-issues/Genome-Editing/happening-right-now)

Questions and answers about CRISPR (https://www.broadinstitute.org/what-broad/areas-focus/project-spotlight/questions-and-answers-about-crispr) are available from the Broad Institute.

The Personal Genetics Education Project has a fact sheet, Genetic Modification, Genome Editing, and CRISPR (https://pged.org/genetic-modification-genome-editing-and-crispr/), that provides an introduction to genome editing.

Yourgenome.org (from the Wellcome Genome Campus) provides information for the public about CRISPR-Cas9 (https://www.yourgenome.org/facts/what-is-crispr-cas9).

A video illustrating how CRISPR-Cas9 works (https://www.youtube.com/watch?v=2pp17E4E-O8) is available from the McGovern Institute for Brain Research at MIT.

ClinicalTrials.gov has a list of human studies using genome editing (https://clinicaltrials.gov/ct2/results?cond=&term=CRISPR+OR+genome+editing+OR+gene+editing) related to various diseases.